CS474 Natural Language Processing

- Last class
  - Intro to lexical semantics
- Today
  - Lexical semantic resources: WordNet
  - Word sense disambiguation
    - » Dictionary-based approaches
    - » Supervised machine learning methods
- New classroom!!!! Hollister 206

CS Colloquium of Possible Interest

Synonymy

- Lexemes with the same meaning
- Invoke the notion of **substitutability**
  - Two lexemes will be considered synonyms if they can be substituted for one another in a sentence without changing the meaning or acceptability of the sentence
    - » How *big* is that plane?
    - » Would I be flying on a *large* or small plane?
    - » Miss Nelson, for instance, became a kind of *big* sister to Mrs. Van Tassel's son, Benjamin.
    - » We frustrate 'em and frustrate 'em, and pretty soon they make a *big* mistake.
    - » Also issues of **register**
      - ◆Social factors that surround the use of possible synonyms, e.g. politeness, group status.

WordNet

- Handcrafted database of lexical relations
- Three separate databases: nouns; verbs; adjectives and adverbs
- Each database is a set of lexical entries (according to unique orthographic forms)
  - Set of senses associated with each entry

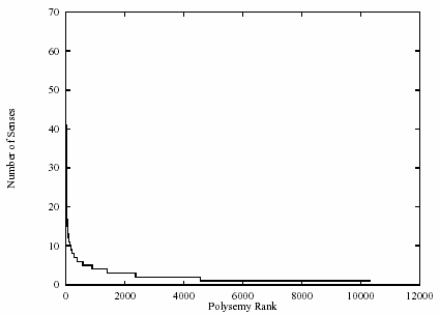
Category	Unique Forms	Number of Senses
Noun	94474	116317
Verb	10319	22066
Adjective	20170	29881
Adverb	4546	5677

Sample entry

The noun "bass" has 8 senses in WordNet.
1. bass - (the lowest part of the musical range)
2. bass, bass part - (the lowest part in polyphonic music)
3. bass, basso - (an adult male singer with the lowest voice)
4. sea bass, bass - (flesh of lean-fleshed saltwater fish of the family Serranidae)
5. freshwater bass, bass - (any of various North American lean-fleshed freshwater fishes especially of the genus Micropterus)
6. bass, bass voice, basso - (the lowest adult male singing voice)
7. bass - (the member with the lowest range of a family of musical instruments)
8. bass - (nontechnical name for any of numerous edible marine and freshwater spiny-finned fishes)

Distribution of senses

- Zipf distribution of senses



WordNet relations

- Nouns

Relation	Definition	Example
Hypernym	From concepts to superordinates	*breakfast → meal*
Hyponym	From concepts to subtypes	*meal → lunch*
Has-Member	From groups to their members	*faculty → professor*
Member-Of	From members to their groups	*copilot → crew*
Has-Part	From wholes to parts	*table → leg*
Part-Of	From parts to wholes	*course → meal*
Antonym	Opposites	*leader → follower*

- Verbs

Relation	Definition	Example
Hypernym	From events to superordinate events	*fly → travel*
Troponym	From events to their subtypes	*walk → stroll*
Entails	From events to the events they entail	*snore → sleep*
Antonym	Opposites	*increase ⟺ decrease*

- Adjectives/adverbs

Relation	Definition	Example
Antonym	Opposite	*heavy ⟺ light*
Adverb	Opposite	*quickly ⟺ slowly*

CS674 Natural Language Processing

- Last class
  - Intro to lexical semantics
- Today
  - Lexical semantic resources: WordNet
  - Word sense disambiguation
    - » Dictionary-based approaches
    - » Supervised machine learning methods
    - » Issues for WSD evaluation

Word sense disambiguation

- Given a *fixed* set of senses associated with a lexical item, determine which of them applies to a particular instance of the lexical item
- Two fundamental approaches
  - WSD occurs during semantic analysis as a side-effect of the elimination of ill-formed semantic representations
  - Stand-alone approach
    - » WSD is performed independent of, and prior to, compositional semantic analysis
    - » Makes minimal assumptions about what information will be available from other NLP processes
    - » Applicable in large-scale practical applications

Dictionary-based approaches

- Rely on machine readable dictionaries
- Initial implementation of this kind of approach is due to Michael Lesk (1986)
  - Given a word *W* to be disambiguated
    - » Retrieve all of the sense definitions, *S*, for *W* from the MRD
    - » Compare each *s* in *S* to the dictionary definitions of all the remaining words in the context
    - » Select the sense *s* with the most overlap with (the definitions of) these context words

Example

- Word: *cone*
- Context: *pine cone*
- Sense definitions

pine 1 kind of evergreen tree with needle-shaped leaves
2 waste away through sorrow or illness

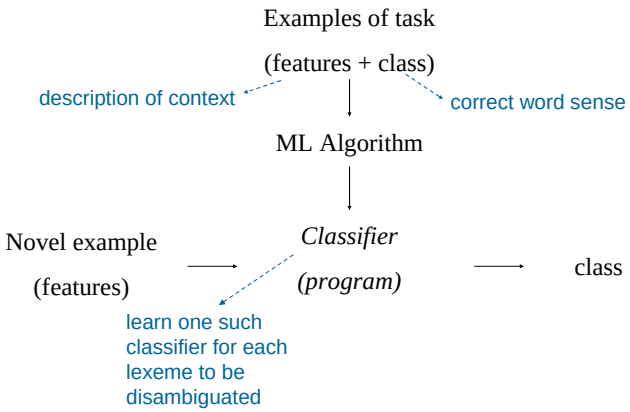
cone 1 solid body which narrows to a point
2 something of this shape whether solid or hollow
3 fruit of certain evergreen trees

- Accuracy of 50-70% on short samples of text from *Pride and Prejudice* and an AP newswire article.

Machine learning approaches

- Machine learning methods
  - Supervised inductive learning
  - Bootstrapping
  - Unsupervised
- Emphasis is on acquiring the knowledge needed for the task from data, rather than from human analysts.

Inductive ML framework



Running example

*An electric guitar and **bass** player stand off to one side, not really part of the scene, just as a sort of nod to gringo expectations perhaps.*

1 Fish sense

2 Musical sense

3 …

Feature vector representation

- **target:** the word to be disambiguated
- **context** : portion of the surrounding text
  - Select a "window" size
  - Tagged with part-of-speech information
  - Stemming or morphological processing
  - Possibly some partial parsing
- Convert the context (and target) into a set of features
  - Attribute-value pairs
    - » Numeric or nominal values

Collocational features

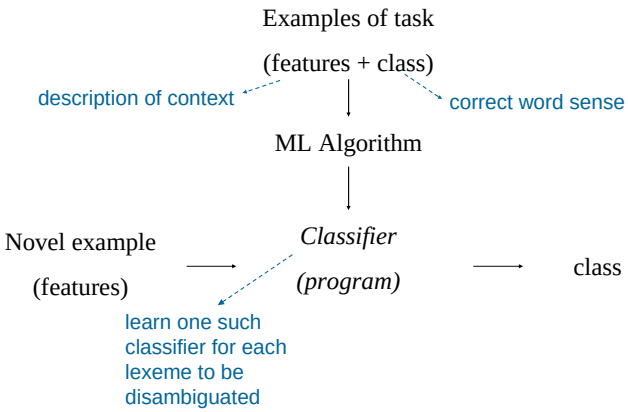
- Encode information about the lexical inhabitants of *specific* positions located to the left or right of the target word.
  - E.g. the word, its root form, its part-of-speech
  - *An electric guitar and **bass** player stand off to one side, not really part of the scene, just as a sort of nod to gringo expectations perhaps.*
  - [guitar, NN1, and, CJC, player, NN1, stand, VVB]

Co-occurrence features

- Encodes information about neighboring words, ignoring exact positions.
  - **Attributes**: the words themselves (or their roots)
  - **Values**: number of times the word occurs in a region surrounding the target word
  - Select a small number of frequently used content words for use as features
    - » 12 most frequent content words from a collection of *bass* sentences drawn from the WSJ: *fishing, big, sound, player, fly, rod, pound, double, runs, playing, guitar, band*
    - » Co-occurrence vector (window of size 10) for the previous example:

      [0,0,0,1,0,0,0,0,0,0,1,0]

Inductive ML framework



Decision list classifiers

- Decision lists: equivalent to simple case statements.
  - Classifier consists of a sequence of tests (usually on a single feature) to be applied to each input example/vector; returns a word sense.
- Continue only until the first applicable test.
- Default test returns the majority sense.

Decision list example

- Binary decision: fish *bass* vs. musical *bass*

Rule		Sense
fish within window	⇒	**bass**^{1}
striped bass	⇒	**bass**^{1}
guitar within window	⇒	**bass**^{2}
bass player	⇒	**bass**^{2}
piano within window	⇒	**bass**^{2}
tenor within window	⇒	**bass**^{2}
sea bass	⇒	**bass**^{1}
play/V *bass*	⇒	**bass**^{2}
river within window	⇒	**bass**^{1}
violin within window	⇒	**bass**^{2}
salmon within window	⇒	**bass**^{1}
on bass	⇒	**bass**^{2}
bass are	⇒	**bass**^{1}

Learning decision lists

- Consists of *generating* and *ordering* individual tests based on the characteristics of the training data
- Generation: every feature-value pair constitutes a test
- Ordering: based on accuracy on the training set

$$abs\left(\log\frac{P(Sense_1 \mid f_i = v_j)}{P(Sense_2 \mid f_i = v_j)}\right)$$

- Associate the appropriate sense with each test